

Deriving Grover’s lower bound from simple physical principles

Ciarán M. Lee^{1,*} and John H. Selby^{1,2,†}

¹*University of Oxford, Department of Computer Science,
Wolfson Building, Parks Road, Oxford OX1 3QD, UK.*

²*Imperial College London, London SW7 2AZ, UK.*

Grover’s algorithm constitutes the optimal quantum solution to the search problem and provides a quadratic speed-up over all possible classical search algorithms. Quantum interference between computational paths has been posited as a key resource behind this computational speed-up. However there is a limit to this interference, at most pairs of paths can ever interact in a fundamental way. Could more interference imply more computational power? Sorkin has defined a hierarchy of possible interference behaviours—currently under experimental investigation—where classical theory is at the first level of the hierarchy and quantum theory belongs to the second. Informally, the order in the hierarchy corresponds to the number of paths that have an irreducible interaction in a multi-slit experiment. In this work, we consider how Grover’s speed-up depends on the order of interference in a theory. Surprisingly, we show that the quadratic lower bound holds regardless of the order of interference. Thus, at least from the point of view of the search problem, post-quantum interference does not imply a computational speed-up over quantum theory.

Grover’s algorithm [12] provides the optimal quantum solution to the search problem and is one of the most versatile and influential quantum algorithms. The search problem—in its simplest form—asks one to find a single “marked” item from an unstructured list of N elements by querying an oracle which can recognise the marked item. The importance of Grover’s algorithm stems from the ubiquitous nature of the search problem and its relation to solving NP-complete problems [6]. Classical computers require $O(N)$ queries to solve this problem, but quantum computers—using Grover’s algorithm—only require $O(\sqrt{N})$ queries. Quantum interference between computational paths has been posited [32] as a key resource behind this computational “speed-up”. However, as first noted by Sorkin [29, 30], there is a limit to this interference—at most pairs of paths can ever interact in a fundamental way. Could more interference imply more computational power?

Sorkin has defined a hierarchy of possible interference behaviours—currently under experimental investigation [24, 27, 28]—where classical theory is at the first level of the hierarchy and quantum theory belongs to the second. Informally, the order in the hierarchy corresponds to the number of paths that have an irreducible interaction in a multi-slit experiment. To get a greater understanding of the role of interference in computation, we consider how Grover’s speed-up depends on the order of interference in a theory.

Restriction to the second level of this hierarchy implies many “quantum-like” features, which, at first glance, appear to be unrelated to interference. For example, such interference behaviour restricts correlations [11] to the “almost quantum correlations” discussed in [21], and bounds contextuality in a manner similar to quantum

theory [14, 23]. This, in conjunction with interference being a key resource in the quantum speed-up, suggests that post-quantum interference may allow for a speed-up over quantum computation.

Surprisingly, we show that this is not the case—at least from the point of view of the search problem. We consider this problem within the framework of generalised probabilistic theories, which is suitable for describing arbitrary operationally-defined theories [5, 8, 9, 13, 16, 17]. Classical probability theory, quantum theory, Spekken’s toy model [15, 31], and the theory of PR boxes [25] all provide examples of theories in this framework. We consider theories satisfying certain natural physical principles which are sufficient for the existence of a well-defined search oracle. Given these physical principles, we prove that a theory at level h in Sorkin’s hierarchy requires $\Omega(\sqrt{N/h})$ queries to solve the search problem. Thus, post-quantum interference does not imply a computational speed-up over quantum theory. Moreover, from the point of view of the search problem, all (finite) orders of interference are asymptotically equivalent.

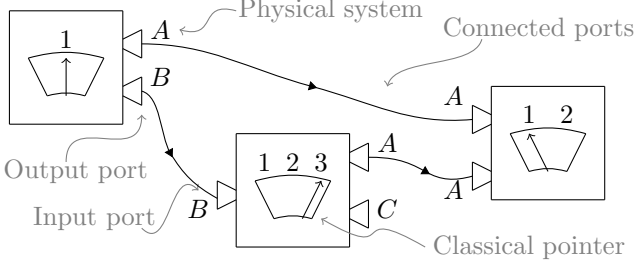
I. GENERALISED PROBABILISTIC THEORIES

A basic requirement of any physical theory is that it should provide a consistent account of experimental data. This idea underlies the framework of generalised probabilistic theories—developed in [4, 8, 9, 13]—which allows for the description of arbitrary theories satisfying this requirement. Informally, a theory in this framework specifies a set of *physical processes* which can be connected together to form experiments. Each process corresponds to a single use of a piece of laboratory apparatus, each having a number of input and output ports, as well as a classical pointer. When the physical apparatus is used in an experiment, the classical pointer comes to rest at one of a number of positions, indicating a specific outcome has occurred. Each port is associated with a *physical sys-*

*Electronic address: ciarane.lee@cs.ox.ac.uk

†Electronic address: john.selby08@ic.ac.uk

tem of a particular type (labelled A, B, \dots). Intuitively one can consider these physical systems as passing from outputs of one process to inputs of another. Processes can thus be connected together—both in sequence and in parallel—to form *circuits*, where it is required that types match and there are no cycles.



Closed circuits (i.e. circuits with no disconnected ports) correspond to the probability of obtaining a particular set of outcomes from the experiment represented by that circuit. Processes that yield the same probabilities in all closed circuits are identified, giving rise to equivalence classes of processes. Each element of such an equivalence class has the same input and output ports, and are denoted ${}_A T_B \in {}_A \mathcal{T}_B$, where ${}_A \mathcal{T}_B$ is the set of possible transformations from systems A to B . Transformations with no input ports are called *states* $S_A \in \mathcal{S}_A$, and no output ports, *effects*, ${}_A E \in {}_A \mathcal{E}$.

Given the probabilistic structure provided by closed circuits, each transformation ${}_A T_B$ can be associated with a real vector such that the set ${}_A \mathcal{T}_B$ is a subset of some real vector space, denoted ${}_A V_B$ [8]. We assume in this work that all vector spaces are finite dimensional. It can be shown that transformations and effects act linearly on the vector space of states, V_A [8]. A measurement corresponds to a set of effects $\{e^r\}$ labelled by the position of the classical pointer r . The probability of preparing state s and observing outcome r is (suppressing system types for readability) given by:

$$e^r(s) = P(r, s).$$

A state is *pure* if it does not arise as a *coarse-graining* of other states [36]; a pure state is one for which we have maximal information. A state is *mixed* if it is not pure. Similarly, one says a transformation is pure if it does not arise as a coarse-graining of other transformations. It can be shown that reversible transformations preserve pure states [9].

We now introduce five physical principles which will be assumed throughout the rest of this work. These can be thought of as an abstraction of basic characteristics of the behaviour of information in quantum theory. Note however that these principles are not unique to quantum theory, indeed, real vector space quantum theory, fermionic quantum theory and the classical theory of pure states each satisfy all of these principles.

Principle 1. Causality [8]: *There exists a unique deterministic effect ${}_A U$ for every system A , such that $\sum_r e^r = U$ for all measurements, $\{e^r\}_r$.*

In quantum theory the unique deterministic effect is provided by the partial trace. Mathematically, causality is equivalent to the statement: “probabilities of present experiments are independent of future measurement choice” [8], and so this can be interpreted as saying that “information propagates from present to future”.

The deterministic effect allows one to define a notion of *marginalisation* for multipartite states.

Principle 2. Purification [8]: *Given a state s_A there exists a system B and a pure state S_{AB} on AB such that s_A is the marginalisation of S_{AB} :*

$${}_B U(S_{AB}) = s_A.$$

Moreover, the purification S_{AB} is unique up to reversible transformations on the purifying system, B [37].

For example, in quantum theory any mixed state $\rho = \sum_i p_i |\psi_i\rangle\langle\psi_i|$ can be written as $\rho = \text{tr}_B(|\Psi\rangle\langle\Psi|_{AB})$ where $|\Psi\rangle_{AB} := \sum_i \sqrt{p_i} |\psi_i\rangle |i\rangle$. Moreover, any other purification $|\tilde{\Psi}\rangle_{AB}$ must satisfy $|\Psi\rangle_{AB} = (\mathbb{I}_A \otimes U_B) |\tilde{\Psi}\rangle_{AB}$ with U_B a unitary transformation. More generally this can be thought of as saying that information cannot be fundamentally destroyed, only discarded.

Principle 3. Purity Preservation [10]: *The composition of pure transformations is pure.*

Pure transformations in quantum theory can be characterised by having Kraus rank 1. Given two such transformations, their sequential or parallel composition will each also be rank 1, and so composition preserves purity.

Principle 4. Pure Sharp Effect [10]: *For each system A there exists a pure effect that occurs with unit probability on some state.*

Pure states $\{a^i\}_{i=1}^n$ are *perfectly distinguishable* if there exists a measurement, corresponding to effects $\{e^j\}_{j=1}^n$, such that $e^j(a^i) = \delta_{ij}$ for all i, j . For example, in quantum theory the computational basis $\{|i\rangle\}$ provide a perfectly distinguishable set, where the corresponding effects are just $\{\langle j|\}$ such that $\langle j|i\rangle = \delta_{ij}$. Such an n -tuple of states can reliably encode an n -level classical system.

Principle 5. Strong symmetry [3]: *For any two n -tuples of pure and perfectly distinguishable states $\{a^i\}$, and $\{b^i\}$, there exists a reversible transformation T such that $T(a^i) = b^i$ for all i .*

An example in quantum theory is the Hadamard transformation reversibly mapping between the bases $\{|0\rangle, |1\rangle\}$ and $\{|+\rangle, |-\rangle\}$.

These last two principles imply that one can encode classical data in a system, and moreover, that any encoding is equivalent. In other words, information is independent of the encoding medium.

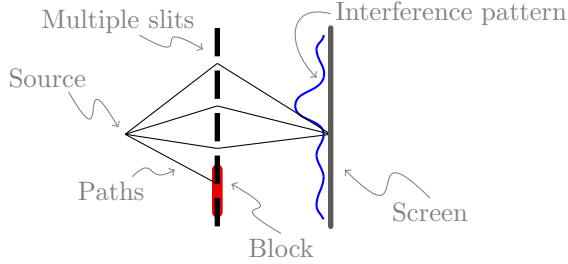
Principles 1 to 4 imply the following result (see [10] for a proof): for any given state s , there exists a natural number n and a set of pure and perfectly distinguishable

states $\{a^i\}_{i=1}^n$ such that $s = \sum_i p_i a_i$ where $0 \leq p_i \leq 1$, $\forall i$ and $\sum_i p_i = 1$.

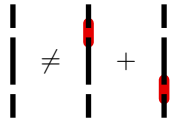
This result, together with principle 5, implies the existence of a “self-dualising” [3, 20] inner product $\langle \cdot, \cdot \rangle$. That is, to every pure state s , there is associated a unique pure effect e^s , satisfying $e^s(s) = 1$, such that: $e^s(\cdot) = \langle s, \cdot \rangle$. This inner product is invariant under all reversible transformations; satisfies $0 \leq \langle r, s \rangle \leq 1$ for all states r, s ; $\langle s, s \rangle = 1$ for all pure states s ; and $\langle s, r \rangle = 0$ if s and r are perfectly distinguishable. It also gives rise to the norm $\| \cdot \| = \sqrt{\langle \cdot, \cdot \rangle}$, satisfying $\|s\| \leq 1$ for all states s , with equality for pure states. We will make use of this norm in proving our main result.

II. HIGHER-ORDER INTERFERENCE

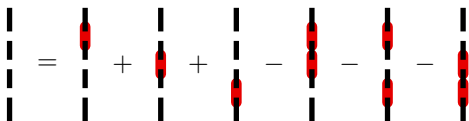
Informally, a theory is said to have n th order interference if one can generate interference patterns in an n -slit experiment which cannot be created in any experiment with only m -slits, for all $m < n$.



More precisely, this means that the interference pattern created on the screen cannot be written as a particular linear combination of the patterns generated when different subsets of slits are blocked. In the two slit experiment, quantum interference corresponds to the fact that the interference pattern cannot be written as the sum of the single slit patterns:



It was first shown by Sorkin [29, 30] that—at least for ideal experiments [26]—quantum theory is limited to the $n = 2$ case. That is, the interference pattern created in a three—or more—slit experiment *can* be written in terms of the two and one slit interference patterns obtained by blocking some of the slits. Schematically:



If a theory does not have n th order interference then one can show it will not have m th order interference, for any $m > n$ [29]. As such, one can classify theories according

to their maximal order of interference, h . For example quantum theory lies at $h = 2$ and classical theory at $h = 1$.

Higher order interference was initially formalised by Sorkin in the framework of Quantum Measure Theory [29] but has more recently been adapted to the setting of generalised probabilistic theories in [3, 18, 19, 33]. The most direct translation to this setting describes the order of interference in terms of probability distributions corresponding to the different experimental setups (which slits are open, etc.) [18]. However, given our five principles, it is possible to define physical transformations that correspond to the action of blocking certain subsets of slits. In this case, there is a more convenient (and equivalent, given the five principles) definition in terms of such transformations [3].

If there are N slits, labelled $1, \dots, N$, these transformations are denoted P_I , where $I \subseteq \{1, \dots, N\} := \mathbf{N}$ corresponds to the subset of slits which are not blocked. In general we expect that $P_I P_J = P_{I \cap J}$, as only those slits belonging to both I and J will not be blocked by either P_I or P_J . This intuition suggests that these transformations should correspond to projectors (i.e. idempotent transformations $P_I P_I = P_I$). Given principles 1 to 5, it was shown in [3] that this is indeed the case. Given this structure, one can define the maximal order of interference as follows [3].

Definition 1. A theory satisfying principles 1 to 5 has maximal order of interference h if, for any $N \geq h$, one has:

$$\mathbb{1}_N = \sum_{\substack{I \subseteq \mathbf{N} \\ |I| \leq h}} \mathcal{C}(h, |I|, N) P_I$$

where $\mathbb{1}_N$ is the identity on a system with N pure and perfectly distinguishable states and

$$\mathcal{C}(h, |I|, N) := (-1)^{h-|I|} \binom{N-|I|-1}{h-|I|}$$

The factor $\mathcal{C}(h, |I|, N)$ in the above definition corrects for the overlaps that occur when different combinations of slits are blocked. Note that, for the case $h = N$, this reduces to the expected expression of $\mathbb{1}_h = P_{\{1, \dots, h\}}$ i.e. the identity is given by the projector with all slits open. The case of $N = h + 1$ corresponds to $\mathcal{C}(h, |I|, h + 1) = (-1)^{h-|I|}$, which is the situation depicted in the previous figures, as well as the one most commonly discussed in the literature [29, 33].

Rather than work directly with these physical projectors, it is mathematically more convenient to work with (generally) unphysical transformations corresponding to projectors onto the “coherences” of a state. For example, in the case of a qutrit, the projector $P_{\{0,1\}}$ projects onto a two dimensional subspace:

$$P_{\{0,1\}} :: \begin{pmatrix} \rho_{00} & \rho_{01} & \rho_{02} \\ \rho_{10} & \rho_{11} & \rho_{12} \\ \rho_{20} & \rho_{21} & \rho_{22} \end{pmatrix} \mapsto \begin{pmatrix} \rho_{00} & \rho_{01} & 0 \\ \rho_{10} & \rho_{11} & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

whilst the coherence-projector $\omega_{\{0,1\}}$ projects only onto the coherences in that two dimensional subspace:

$$\omega_{\{0,1\}} :: \begin{pmatrix} \rho_{00} & \rho_{01} & \rho_{02} \\ \rho_{10} & \rho_{11} & \rho_{12} \\ \rho_{20} & \rho_{21} & \rho_{22} \end{pmatrix} \mapsto \begin{pmatrix} 0 & \rho_{01} & 0 \\ \rho_{10} & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

That is, $\omega_{\{0,1\}}$ corresponds to the linear combination of projectors: $P_{\{0,1\}} - P_{\{0\}} - P_{\{1\}}$.

There is a coherence-projector ω_I for each subset of slits $I \subseteq \mathbf{N}$, defined in terms of the physical projectors:

$$\omega_I := \sum_{\tilde{I} \subseteq I} (-1)^{|\tilde{I}|+|I|} P_{\tilde{I}}.$$

These have the following useful properties, proved in appendix A.

Lemma 1. *An equivalent definition of the maximal order of interference, h , is: $\mathbb{1}_N = \sum_{I, |I|=1}^h \omega_I$, for all $N \geq h$.*

The above lemma implies that any state (indeed, any vector in the vector space generated by the states) can be decomposed as $s = \sum_{I, |I|=1}^h s_I$, where $s_I := \omega_I s$.

Lemma 2. *“Coherences are orthogonal”: i) $\omega_I \omega_J = \delta_{IJ} \omega_I$, for all I, J and ii) $\|s\|^2 = \sum_I \|\omega_I s\|^2$*

III. SETTING UP THE PROBLEM

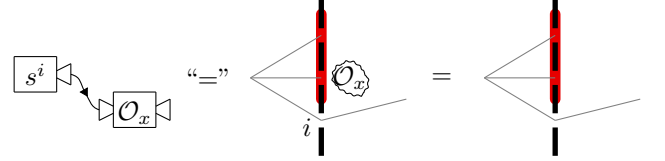
In the standard search problem, one is asked to find a specific “marked” item from among a large collection of items in some unstructured list. The items are indexed $1, \dots, N$ and one has access to an oracle, which, when asked whether item i is the marked item, denoted x , returns the answer “yes” or “no”. Informally, the search problem asks for the minimal number of queries to this oracle required to find x in the worst case.

In the standard bra-ket formalism of quantum theory, this oracle corresponds to a controlled unitary transformation U , defined by its action on the (product) computational basis: $U|i\rangle|q\rangle = |i\rangle|q \oplus f(i)\rangle$, where $|i\rangle$ is the index, or control, register, $|q\rangle$ is the target register, \oplus denotes addition modulo 2 and $f : \{1, \dots, N\} \rightarrow \{0, 1\}$ satisfies $f(i) = 1$ if and only if $i = x$. Inputting $|-\rangle$ into the target register results in a phase being “kicked-back” to the control register: $U|i\rangle|-\rangle = (-1)^{f(i)}|i\rangle|-\rangle$. Discarding the target register reduces the action of the oracle to applying the phase transformation $\mathcal{O}_x|i\rangle = (-1)^{f(i)}|i\rangle$. Changing to the density matrix formalism, we see that this phase oracle, whose action on states ρ is now denoted by $\mathcal{O}_x \rho$, acts as the identity on the diagonal elements of all density matrices whilst adding a ‘-’ to the off diagonal elements $\{\rho_{xi}, \rho_{ix}\}_i$.

Previous work [18] has shown that the conjunction of principles 1, 2, 3 and 5 implies the existence of reversible controlled transformations. These can be used to define oracles in a manner analogous to quantum theory [18].

Moreover, every controlled transformation gives rise to a “kicked-back” reversible phase transformation on the control system [18]. Thus—as in quantum theory—from the point of view of querying the oracle, we can reduce all considerations involving the controlled transformation to those involving the kicked-back phase.

To highlight the role of interference in searching an unstructured list, we describe the action of querying the oracle in terms of the physically motivated set-up of N -slit experiments. Consider first the quantum case. Note that an N -slit experiment defines a set of N pure and perfectly distinguishable states $|i\rangle\langle i|$, each of which can be associated to a distinct element in the N item list. Querying the oracle about item i is equivalent to applying the oracle transformation to state $|i\rangle\langle i|$. In quantum theory, preparing such a state can be achieved by passing a uniform superposition through the N -slit experiment with all but the i th slit blocked. The oracle can be implemented by placing a phase shifter behind slit x . Querying the oracle in a superposition of states can then be achieved by varying which slits are blocked. This is illustrated schematically below:



As discussed previously, the physical act of blocking slits is represented by the projectors P_I . The action of the quantum oracle can thus be rephrased in terms of these projectors: i) $\mathcal{O}_x P_I = P_I$, if $x \notin I$ or $|I| = 1$ and, ii) \mathcal{O}_x can act non-trivially on projectors P_I with $x \in I$ and $|I| > 1$, but must satisfy $\mathcal{O}_x P_I = P_I \mathcal{O}_x$, for all P_I , which corresponds to the fact that a quantum oracle does not “create” or “destroy” coherence between states passing through different slits.

By analogy with the quantum case we can define the oracle which encodes the search problem in theories satisfying principles 1 to 5 as follows. Note that in this paper we only deal with the case of a single marked item.

Definition 2. *A reversible transformation is a search oracle, denoted \mathcal{O}_x , if and only if:*

- i) $\mathcal{O}_x P_I = P_I$ for all $x \notin I$ or $|I| = 1$ and,
- ii) $\mathcal{O}_x P_I = P_I \mathcal{O}_x$, for all P_I .

In the above definition, the requirement $\mathcal{O}_x P_I = P_I \mathcal{O}_x$, for all P_I , is quite natural. This requirement ensures that one cannot gain any information about item i when querying the oracle using a state with no support on i , i.e. a state s such that $P_I s = s$ where $i \notin I$. In an arbitrary theory, it may not be the case that a transformation satisfying definition 2 and acting non-trivially on P_I , with $x \in I$, exists. This is not an issue as in such theories we cannot even define the search problem, let alone show it can be solved using fewer queries than

quantum theory. In this work, we shall assume the existence of a search oracle in any theory we consider. Given the definition of coherence-projectors ω_I we can equivalently write definition 2 as: $\mathcal{O}_x \omega_I = \omega_I$, for $x \notin I$ or $|I| = 1$, and $\mathcal{O}_x \omega_I = \omega_I \mathcal{O}_x$, for all I . Indeed, in the quantum case, the action of the oracle can be equivalently described as: $\mathcal{O}_x \omega_I = \omega_I$ if $x \notin I$ or $|I| = 1$, and $\mathcal{O}_x \omega_I = -\omega_I$ otherwise.

We can now formally state the search problem for a single marked item—defined for the quantum case in [7, 22, 34]—as:

Search Problem. *Given an N element list with search oracle \mathcal{O}_x and an arbitrary collection of reversible transformations $\{G_i\}$, what is the minimal $k \in \mathbb{N}$ such that $G_k \mathcal{O}_x G_{k-1} \dots G_1 \mathcal{O}_x s$ can be found, with probability greater than $1/2$, to be in the state x , for arbitrary state s , averaged over all possible marked items?*

IV. MAIN RESULT

Theorem 1. *In theories satisfying principles 1 to 5, with finite maximal order of interference h , the number of queries needed to solve the search problem is $\Omega(\sqrt{N/h})$.*

Proof of theorem 1. The basic idea is based on the proof of the quantum case presented in [7, 22, 34]. Let

$$\begin{aligned} s_k^x &= G_k \mathcal{O}_x G_{k-1} \dots G_1 \mathcal{O}_x s, \\ s_k &= G_k G_{k-1} \dots G_1 s, \end{aligned}$$

where G_i is some reversible transformation from the theory, and define

$$D_k = \sum_x \|s_k^x - s_k\|^2.$$

It will be shown that, for $\langle x, s_k^x \rangle \geq 1/2$, we have $cN \leq D_k \leq 4hk^2$, where c is any constant less than $(\sqrt{2} - 1)^2$, from which the result $k \geq O\left(\sqrt{\frac{N}{h}}\right)$ follows. The lower bound goes through as in the quantum case and is derived in appendix A 4. The upper bound will now be proved by induction.

We have

$$\begin{aligned} D_{k+1} &= \sum_x \|G_{k+1}(\mathcal{O}_x s_k^x - s_k)\|^2 = \sum_x \|\mathcal{O}_x s_k^x - s_k\|^2 \\ &= \sum_x \|\mathcal{O}_x(s_k^x - s_k) + (\mathcal{O}_x - \mathbb{1})s_k\|^2 \\ &\leq \sum_x \|s_k^x - s_k\|^2 \\ &\quad + 2 \sum_x \|\mathcal{O}_x(s_k^x - s_k)\| \|(\mathcal{O}_x - \mathbb{1})s_k\| \\ &\quad + \sum_x \|(\mathcal{O}_x - \mathbb{1})s_k\|^2 \\ &\leq D_k + 2 \sqrt{D_k \sum_x \|(\mathcal{O}_x - \mathbb{1})s_k\|^2} + \|(\mathcal{O}_x - \mathbb{1})s_k\|^2 \\ &\leq \left(\sqrt{D_k} + \sqrt{\sum_x \|(\mathbb{1} - \mathcal{O}_x)s_k\|^2} \right)^2, \end{aligned}$$

which follows from the triangle inequality, the Cauchy-Schwarz inequality, and the fact the norm is invariant under reversible transformations.

The quantity $\sum_x \|(\mathbb{1} - \mathcal{O}_x)s_k\|^2$ —which can be thought of as how much some state is “moved” in a single query, averaged over all possible marked items x —is the only theory dependent quantity that features in this proof. We upper bound it as follows:

$$\begin{aligned} &\sum_x \|(\mathbb{1} - \mathcal{O}_x)s_k\|^2 \\ &= \sum_x \sum_I \|(\mathbb{1} - \mathcal{O}_x)\omega_I s_k\|^2 \\ &= \sum_x \sum_{\substack{I \\ |I| > 1 \\ x \in I}} \|\omega_I(\mathbb{1} - \mathcal{O}_x)s_k\|^2 \\ &\leq \sum_x \sum_{\substack{I \\ |I| > 1 \\ x \in I}} (\|\mathbb{1}\omega_I s_k\| + \|\mathcal{O}_x \omega_I s_k\|)^2 \\ &\leq \sum_x \sum_{\substack{I \\ |I| > 1 \\ x \in I}} 4\|\omega_I s_k\|^2, \end{aligned}$$

where the first line follows from lemma 1, lemma 2, and the definition of the search oracle \mathcal{O}_x , and second from the triangle inequality and the fact that the norm is invariant under reversible transformations. We need to know how many times each $\|\omega_I s_k\|^2$ appears when we sum over the marked item x . Each given $I = \{i_1, i_2, \dots, i_{|I|}\}$ will appear $|I|$ times as we sum over x , one for every time i_j is the marked item. Thus

$$\begin{aligned} \sum_x \|(\mathbb{1} - \mathcal{O}_x)s_k\|^2 &\leq \sum_{\substack{I \\ |I| > 1}} 4|I| \|\omega_I s_k\|^2 \\ &\leq 4 \sum_I |I| \|\omega_I s_k\|^2 \leq 4h \sum_I \|\omega_I s_k\|^2 = 4h \|s_k\|^2 \leq 4h. \end{aligned}$$

The second line follows from $\sum_{|I|=1} \|\omega_I s_k\|^2 \geq 0$, lemma 2, $\|s_k\| \leq 1$, and $|I| \leq h$, for all I . We thus have: $D_{k+1} \leq \left(\sqrt{D_k} + \sqrt{4h}\right)^2$. Assuming that $D_k \leq 4hk^2$ gives us $D_{k+1} \leq 4h(k+1)^2$, from which the result follows via induction. \square

V. DISCUSSION

In this work, we considered theories satisfying certain natural physical principles which are sufficient for the existence of controlled transformations and a phase kick-back mechanism, necessary features for a well-defined search oracle. Given these physical principles, we proved that a theory with maximal order of interference h requires $\Omega(\sqrt{N/h})$ queries to this oracle to find a single marked item from some N -element list. This result challenges our pre-conceived notions about how quantum computers achieve their computational advantage and is somewhat surprising as one might expect more interference to imply more computational power. Further work will focus on determining sufficient physical principles for there to exist an algorithm that achieves the quadratic lower bound derived here.

Recent work has also investigated Grover's algorithm from the point of view of post-quantum theories [1, 2]. These works considered modifications of quantum theory which allow for superluminal signalling and cloning of states. In contrast, the generalised probabilistic theory framework employed here allowed us to investigate Grover's lower bound in alternate theories that are physically reasonable and which, for example, do not allow for superluminal signalling [4] or cloning [35].

As theories satisfying our five physical principles appear 'quantum-like'—at least from the point of view of the search problem—investigating interference behaviour in them may inform current experiments searching for post-quantum interference.

Acknowledgements—The authors thank H. Barnum and M. J. Hoban for useful discussions and M. J. Hoban for proof reading a draft of the current paper. The authors also acknowledge encouragement and support from J. J. Barry. This work was supported by the EPSRC through the Controlled Quantum Dynamics Centre for Doctoral Training and the Oxford Department of Computer Science. CML also acknowledges funding from University College, Oxford.

Appendix A: Results for coherences

1. Proof of lemma 1

In a theory with maximal order of interference h one has

$$\mathbb{1}_N = \sum_{\substack{I \subseteq \mathbf{N} \\ |I| \leq h}} \mathcal{C}(h, |I|, N) P_I.$$

Thus, showing $\mathbb{1}_N = \sum_{|I|=1}^h \omega_I$ reduces to showing

$$\sum_{|I|=1}^h \omega_I = \sum_{\substack{I \subseteq \mathbf{N} \\ |I| \leq h}} \mathcal{C}(h, |I|, N) P_I.$$

As $\omega_I := \sum_{\tilde{I} \subseteq I} (-1)^{|I|+|\tilde{I}|} P_{\tilde{I}}$, we just have to count the number of P_I 's that appear as we sum over $|I|$. For some fixed I , this is just

$$\sum_{\alpha=|I|}^h (-1)^{\alpha-|I|} \binom{N-|I|}{\alpha-|I|}.$$

By expanding and rearranging this, one can straightforwardly (if tediously) show that this equals $\mathcal{C}(h, |I|, N)$, and we are done.

2. Proof of lemma 2 part i)

From the definition of ω_I , it follows that

$$\begin{aligned} \omega_I \omega_J &= (-1)^{|I|+|J|} \sum_{\tilde{I} \subseteq I} \sum_{\tilde{J} \subseteq J} (-1)^{|\tilde{I}|+|\tilde{J}|} P_{\tilde{I}} P_{\tilde{J}} \\ &= (-1)^{|I|+|J|} \sum_{\tilde{K} \subseteq I \cap J} \mathcal{D}(I, J, \tilde{K}) P_{\tilde{K}} \end{aligned}$$

where $\mathcal{D}(I, J, \tilde{K})$ is the number of distinct pairings of \tilde{I} and \tilde{J} such that $\tilde{I} \cap \tilde{J} = \tilde{K}$ and $|\tilde{I}| + |\tilde{J}|$ is even, minus the number of distinct pairings where $\tilde{I} \cap \tilde{J} = \tilde{K}$ and $|\tilde{I}| + |\tilde{J}|$ is odd. It will now be shown that

$$\mathcal{D}(I, J, \tilde{K}) = \begin{cases} 0 & \text{if } I \neq J \\ (-1)^{|I|+|\tilde{K}|} & \text{if } I = J \end{cases}$$

For the $I \neq J$ case fix some particular $i \in I$ such that $i \notin J$ and consider some $\tilde{I} \subseteq I, \tilde{J} \subseteq J$ such that $\tilde{I} \cap \tilde{J} = \tilde{K}$. If $x \notin \tilde{I}$ alter \tilde{I} by adding i , otherwise alter \tilde{I} by removing x . This procedure turns each even $|\tilde{I}| + |\tilde{J}|$, odd. We have thus shown that for each $\tilde{I} \subseteq I$ and $\tilde{J} \subseteq J$ such that $\tilde{I} \cap \tilde{J} = \tilde{K}$ and $|\tilde{I}| + |\tilde{J}|$ is even, there exists

an $\tilde{I}' \subseteq I$ such that $\tilde{I}' \cap \tilde{J} = \tilde{K}$ and $|\tilde{I}'| + |\tilde{J}|$ is odd, and vice versa. Thus the number of distinct pairings of \tilde{I} and \tilde{J} such that $\tilde{I} \cap \tilde{J} = \tilde{K}$ and $|\tilde{I}| + |\tilde{J}|$ is even is equal to the number of distinct pairings of \tilde{I} and \tilde{J} such that $\tilde{I} \cap \tilde{J} = \tilde{K}$ and $|\tilde{I}| + |\tilde{J}|$ is odd, and so $\mathcal{D}(I, J, \tilde{K}) = 0$ when $I \neq J$.

For the $I = J$ case we can make a similar argument by picking some $i \in I, i \notin \tilde{J}$ except for when $\tilde{J} = J = I$. This case gives an excess ± 1 depending on whether $|J| + |\tilde{K}|$ is odd or even, implying $\mathcal{D}(I, J, \tilde{K}) = (-1)^{|I|+|\tilde{K}|}$ when $I = J$.

This immediately gives $\omega_I \omega_J = 0$ if $I \neq J$ and,

$$\omega_I \omega_I = (-1)^{2|I|} \sum_{\tilde{K} \subseteq I} (-1)^{|I|+|\tilde{K}|} P_{\tilde{K}} = \omega_I$$

if $I = J$.

3. Proof of lemma 2 part ii)

To prove the lemma, we need the fact that the ω_I 's are self-dual $\omega_I^\dagger = \omega_I$, where the \dagger is defined by the self-dualising inner-product as: $\langle \cdot, \omega_I \cdot \rangle = \langle \omega_I^\dagger \cdot, \cdot \rangle$. Recalling that the ω_I 's correspond to linear combinations of the P_I 's, this follows immediately from self-duality of the projectors P_I , which is proved in [3] (Recall that principles 1 to 5 imply the first two axioms of [3]). We now have

$$\begin{aligned} \|s\|^2 &= \langle s, s \rangle = \left\langle \sum_I \omega_I s, \sum_J \omega_J s \right\rangle \\ &= \sum_{I,J} \langle \omega_I s, \omega_J s \rangle = \sum_{I,J} \langle s, \omega_I^\dagger \omega_J s \rangle \end{aligned}$$

$$= \sum_{I,J} \langle s, \omega_I \omega_J s \rangle = \sum_{I,J} \delta_{IJ} \langle s, \omega_I s \rangle$$

where the last equality follows from the orthogonality of the ω_I 's. Finally

$$\|s\|^2 = \sum_I \langle s, \omega_I s \rangle = \sum_I \langle \omega_I s, \omega_I s \rangle = \sum_I \|\omega_I s\|^2$$

4. Proof of $D_k \geq cN$

We assume that $\langle x, s_k^x \rangle \geq 1/2$ for all x , so a measurement of s_k^x yields a solution to the search problem with probability at least $1/2$. Let $E_k = \sum_x \|s_k^x - x\|^2$ and $F_k = \sum_x \|s_k - x\|^2$. It follows that

$$\begin{aligned} \text{i) } E_k &= \sum_x 2(1 - \langle x, s_k^x \rangle) \leq \sum_x 2(1 - 1/2) \leq N \text{ and,} \\ \text{ii) } F_k &\geq 2 \left(N - \|s_k\| \sqrt{\left\langle \sum_x x, \sum_y y \right\rangle} \right) \geq 2(N - \sqrt{N}) \end{aligned}$$

where ii) follows from the Cauchy-Schwarz inequality, $\|s_k\| \leq 1$ and $\langle x, y \rangle = \delta_{xy}$. As explicitly calculated on page 270 of [22], by using the reverse triangle inequality and the Cauchy-Schwarz inequality, it follows that $D_k \geq (\sqrt{F_k} - \sqrt{E_k})^2$. Combining this with the upper bound on E_k and the lower bound on F_k , we have that $D_k \geq cN$, for sufficiently large N , where c is any constant less than $(\sqrt{2} - 1)^2 \approx 0.17$.

-
- [1] Scott Aaronson, Adam Bouland, Joseph Fitzsimons, and Mitchell Lee. The space just above bqp. In *Proceedings of the 2016 ACM Conference on Innovations in Theoretical Computer Science*, pages 271–280. ACM, 2016.
 - [2] Ning Bao, Adam Bouland, and Stephen P Jordan. Grover search and the no-signaling principle. *arXiv preprint arXiv:1511.00657*, 2015.
 - [3] Howard Barnum, Markus P Müller, Cozmin Ududec, et al. Higher-order interference and single-system postulates characterizing quantum theory. *New Journal of Physics*, 16(12):123029, 2014.
 - [4] Jonathan Barrett. Information processing in generalized probabilistic theories. *Physical Review A*, 75(3):032304, 2007.
 - [5] Jonathan Barrett, Niel de Beaudrap, Ciarán M Lee, and Matty J Hoban. The computational landscape of general physical theories. *Forthcoming*, 2016.
 - [6] Charles H Bennett, Ethan Bernstein, Gilles Brassard, and Umesh Vazirani. Strengths and weaknesses of quantum computing. *SIAM journal on Computing*, 26(5):1510–1523, 1997.
 - [7] Michel Boyer, Gilles Brassard, Peter Høyer, and Alain Tapp. Tight bounds on quantum searching. *arXiv preprint quant-ph/9605034*, 1996.
 - [8] Giulio Chiribella, Giacomo Mauro D’Ariano, and Paolo Perinotti. Probabilistic theories with purification. *Physical Review A*, 81(6):062348, 2010.
 - [9] Giulio Chiribella, Giacomo Mauro D’Ariano, and Paolo Perinotti. Informational derivation of quantum theory. *Physical Review A*, 84(1):012311, 2011.
 - [10] Giulio Chiribella and Carlo Maria Scandolo. Operational axioms for state diagonalization. *arXiv preprint arXiv:1506.00380*, 2015.
 - [11] Fay Dowker, Joe Henson, and Petros Wallden. A histories perspective on characterizing quantum non-locality. *New Journal of Physics*, 16(3):033033, 2014.
 - [12] Lov K Grover. Quantum mechanics helps in searching for a needle in a haystack. *Physical review letters*, 79(2):325, 1997.
 - [13] Lucien Hardy. Reformulating and reconstructing quan-

- tum theory. *arXiv preprint arXiv:1104.2066*, 2011.
- [14] Joe Henson. Bounding quantum contextuality with lack of third-order interference. *Phys. Rev. Lett.* (*arXiv:1406.3281*), (114):220403, 2015.
 - [15] Peter Janotta and Raymond Lal. Generalized probabilistic theories without the no-restriction hypothesis. *Physical Review A*, 87(5):052131, 2013.
 - [16] Ciarán M Lee and Jonathan Barrett. Computation in generalised probabilistic theories. *New Journal of Physics*, 17(8):083001, 2015.
 - [17] Ciarán M Lee and Matty J Hoban. Proofs and advice in general physical theories: a trade-off between states and dynamics? *arXiv preprint arXiv:1510.04702*, 2015.
 - [18] Ciarán M Lee and John H Selby. Generalised phase kickback: a computational advantage for higher-order interference? *arXiv preprint arXiv:1510.04699*, 2015.
 - [19] Ciarán M Lee and John H Selby. Higher-order interference in extensions of quantum theory. *arXiv preprint arXiv:1510.03860*, 2015.
 - [20] Markus P Müller and Cozmin Ududec. Structure of reversible computation determines the self-duality of quantum theory. *Physical Review Letters*, 108(13):130401, 2012.
 - [21] Miguel Navascués, Yelena Guryanova, Matty J Hoban, and Antonio Acín. Almost quantum correlations. *Nature communications*, 6, 2015.
 - [22] Michael A Nielsen and Isaac L Chuang. *Quantum computation and quantum information*. Cambridge university press, 2010.
 - [23] Gerd Niestegge. Conditional probability, three-slit experiments, and the jordan algebra structure of quantum mechanics. *Advances in Mathematical Physics*, 2012, 2012.
 - [24] Daniel K Park, Osama Moussa, and Raymond Laflamme. Three path interference using nuclear magnetic resonance: a test of the consistency of born’s rule. *New Journal of Physics*, 14(11):113025, 2012.
 - [25] Sandu Popescu and Daniel Rohrlich. *Causality and non-locality as axioms for quantum mechanics*. Springer, 1998.
 - [26] Aninda Sinha, Aravind H Vijay, and Urbasi Sinha. On the superposition principle in interference experiments. *Scientific reports*, 5, 2015.
 - [27] Urbasi Sinha, Christophe Couteau, Thomas Jennewein, Raymond Laflamme, and Gregor Weihs. Ruling out multi-order interference in quantum mechanics. *Science*, 329(5990):418–421, 2010.
 - [28] Urbasi Sinha, Christophe Couteau, Zachari Medendorp, Immo Söllner, Raymond Laflamme, Rafael Sorkin, and Gregor Weihs. Testing born’s rule in quantum mechanics with a triple slit experiment. *arXiv preprint arXiv:0811.2068*, 2008.
 - [29] Rafael D Sorkin. Quantum mechanics as quantum measure theory. *Modern Physics Letters A*, 9(33):3119–3127, 1994.
 - [30] Rafael D Sorkin. Quantum measure theory and its interpretation. *arXiv preprint gr-qc/9507057*, 1995.
 - [31] Robert W Spekkens. Evidence for the epistemic view of quantum states: A toy theory. *Physical Review A*, 75(3):032110, 2007.
 - [32] Dan Stahlke. Quantum interference as a resource for quantum speedup. *Physical Review A*, 90(2):022302, 2014.
 - [33] Cozmin Ududec, Howard Barnum, and Joseph Emerson. Three slit experiments and the structure of quantum theory. *Foundations of Physics*, 41(3):396–405, 2011.
 - [34] Christof Zalka. Grover’s quantum searching algorithm is optimal. *Physical Review A*, 60(4):2746, 1999.
 - [35] Howard Barnum, Jonathan Barrett, Matthew Leifer, and Alexander Wilce. A generalized no-broadcasting theorem. *Phys. Rev. Lett.*, 99:240501, 2007.
 - [36] The process $\{\mathcal{U}_j\}_{j \in Y}$, where j indexes the classical pointer, is a coarse-graining of $\{\mathcal{E}_i\}_{i \in X}$ if there is a disjoint partition $\{X_j\}_{j \in Y}$ of X such that $\mathcal{U}_j = \sum_{i \in X_j} \mathcal{E}_i$.
 - [37] Two states S_{AB} and S'_{AB} purifying s_A satisfy $S_{AB} = \mathbb{I} \otimes_B T_B(S'_{AB})$, with ${}_B T_B$ a reversible transformation.